Cetacean Mitochondrial DNA Control Region: Sequences of All Extant Baleen Whales and Two Sperm Whale Species

Ólafur Árnason, Anette Gullberg, and Bengt Widegren

Department of Molecular Genetics, University of Lund

The sequence of the mitochondrial control region was determined in all 10 extant species commonly assigned to the suborder Mysticeti (baleen or whalebone whales) and to two odontocete (toothed whale) species (the sperm and the pygmy sperm whale). In the mysticetes, both the length and the sequence of the control region were very similar, with differences occurring primarily in the first \( \approx 160 \) bp of the 5' end of the L-strand of the region. There were marked differences between the mysticete and sperm whale sequences and also between the two sperm whales. The control region, less its variable portion, was used in a comparison including the 10 mysticete sequences plus the same region of an Antarctic minke whale specimen and the two sperm whales. The difference between the minke whales from the North Atlantic and the Antarctic was greater than that between any acknowledged species belonging to the same genus (Balaenoptera). The difference was similar to that between the families Balaenopteridae (rorquals) and Eschrichtiidae (gray whales). The findings suggest that the Antarctic minke whale should have a full species status, \( B. \) bonaerensis. Parsimony analysis separated the bowhead and the right whale (family Balaenidae) from all remaining mysticetes, including the pygmy right whale. The pygmy right whale is usually included in family Balaenidae. The analysis revealed a close relationship between the gray whale (family Eschrichtiidae) sequence and those of the rorquals (family Balaenopteridae). The gray whale was included in a clade together with the sei, Bryde's, fin, blue, and humpback whales. This clade was separated from the two minke whale types, which branched together.

Introduction

It is a common notion that the mitochondrial control region constitutes the most rapidly evolving portion of mitochondrial DNA (mtDNA) of mammals (e.g., see Saccone et al. 1991). So far, however, species comparisons of complete control-region sequences have mainly been performed at levels above that of the genus. Recently the entire mtDNA molecule of two rorquals (genus Balaenoptera), the blue whale (\( B. \) musculus), and the fin whale (\( B. \) physalus) was sequenced (Árnason et al. 1991a; Árnason and Gullberg, accepted). Comparison between the two species showed that the total difference between their control regions was similar to that of the rest of the mtDNA molecule. The comparison also showed that the differences occurred primarily within the first 160 bp at the 5' end of the L-strand of the control region.

Hoelzel et al. (1991) compared the control region of the Commerson dolphin (Cephalorhynchus commersonii) (Southern et al. 1988), the killer whale (Orcinus

1. Key words: mitochondrial DNA, control region, cetaceans, whalebone whales, sperm whales.

Address for correspondence and reprints: Ólafur Árnason, Department of Molecular Genetics, The Wallenberg Laboratory, Box 7031, S-220 07, Lund, Sweden.

© 1993 by The University of Chicago. All rights reserved.
0737-4038/93/1000-0003$02.00

960
orca), and the minke whale, (B. acutorostrata) (Antarctic specimen), and Árnason et al. (1991a) compared the same region of the fin whale and the Commerson dolphin. These studies, although limited to one mysticete (whalebone whale) genus and one odontocete (toothed whale) family, the Delphinidae, showed that there was great similarity between these particular representatives of the two cetacean suborders, the Mysticeti and the Odontoceti.

The above findings suggested that the evolution of cetacean mtDNA control region is slow or that the separation between the dolphin and mysticete lineages occurred considerably later than the \( \geq 40 \) Myr before present (MYBP) suggested by paleontological studies on the evolutionary separation of odontocetes and mysticetes (Barnes et al. 1985). This figure implies a monophyletic evolution of extant cetaceans, a fact not universally acknowledged until cytogenetic (Árnason 1969, 1974; Duffield 1977) and molecular evidence at the DNA level (Árnason 1982; Árnason et al. 1984) became available.

A comparison between portions of the mitochondrial 16S rRNA gene of the Commerson dolphin and the fin whale (Árnason et al. 1991a) indicated that the odontocete and mysticete lineages shared ancestry that was more recent than that suggested by paleontological findings. These observations have also been supported by comparisons involving portions of the mtDNA of the sperm whale (Physeter macrocephalus) (Árnason et al. 1993).

Because the mitochondrial control regions of the delphinids and the mysticetes had shown unexpectedly large similarity, we investigated the sequence of this region in all extant mysticete species, in order to determine interspecific differences. The study was complemented with the corresponding sequences of the sperm and the pygmy sperm (Kogia breviceps) whales, included because the sperm whales (family Physeteridae) have a rather separate position among cetaceans, as clearly demonstrated by their karyotypes, which deviate radically from those of all other cetaceans, both odontocetes and mysticetes (Árnason and Benirschke 1973; Árnason 1974, 1981a, 1981b).

The now-reported minke whale specimen was taken in Icelandic waters and belonged to B. acutorostrata acutorostrata. Wada and Numachi (1991) showed that there are conspicuous allozyme differences between minke whales from different areas. In some cases the differences even exceed those between different species of the genus Balaenoptera. The molecular findings of Wada and Numachi (1991) suggested that B. acutorostrata should probably be separated into different species. This particular issue was examined further in the present study by comparing the mitochondrial control region of a North Atlantic minke whale specimen with the sequence presented by Hoelzel et al. (1991). The latter sequence originated from a specimen caught in Antarctic Catch Area V (130°E–170°W) and belonged to the minke whale subspecies B. a. bonaerensis (A. R. Hoelzel, personal communication). Thus the present study made it possible to compare the mitochondrial control regions of two distant types of the minke whale and to evaluate the differences in relation to those that occur between acknowledged species of the genus Balaenoptera.

Material and Methods

The material of the fin, blue, sei (Balaenoptera borealis), minke (B. a. acutorostrata), humpback (Megaptera novaeangliae), and sperm whales originated from Icelandic waters. The humpback and blue whale DNA samples were from cell cultures established from biopsies collected from free-swimming animals. The remaining mysticete samples were the same as those used in a recent study applying molecular hy-
bridizations on all mysticete species (Árnason and Best 1991). The sample of the pygmy sperm whale was provided by Dr. James G. Mead of The Smithsonian Institution, Washington, D.C. The specimen was from the U.S. East Coast.

The sequence of the mitochondrial control region of the fin whale is from a directly cloned portion of the mtDNA molecule (Arnason et al. 1991b). All other sequences were PCR amplified and cloned before sequencing. The sequences of the two PCR primers were 5'-GATCGTGAAGCTTCCCCCTCTAGACTCAAGGAAG-3' and 5'-TAGAATTCTAGACATTTTCAGT-3', respectively. The underlined portion of the first primer is identical to positions 15812-15832 of the L-strand of the mtDNA of the fin whale (Arnason et al. 1991a). A HindIII site was appended to the sequence, for cloning purposes. The primer includes portions of tRNA-Thr and tRNA-Pro. The underlined portion of the second primer is identical to positions 472-454 (opposite orientation) in the mtDNA of the same species. The primer is located within tRNA-Phe. An EcoRI site was added to this sequence. Cloning was carried out in both M13mp18 and M13mp19. Sequencing was according to the dideoxy termination protocol (Sanger 1981) with 35S-adenine and used a collection of specific sequencing primers. Sequences were determined from three clones of each specimen.

Alignment of sequences was performed with the programs LineUp and PileUp of the GCG program package (Devereux et al. 1984), with gap weight 2.5 and gap length weight 0.15. The parsimony analysis was performed with PAUP, version 3.1 (Swofford 1993). The GenBank accession numbers of the presented sequences are X72006 for the minke whale and X72195-X72204 for the other unpublished sequences. Users of the sequences are requested to refer to the present paper rather than to the accession numbers.

**Results**

Figure 1 shows 11 mysticete sequences of the L-strand of the mitochondrial control region. The sequences include the previously published fin whale (Árnason et al. 1991a) and blue whale (Amason et al. 1991b) plus two minke whales, one from the North Atlantic and one from the Antarctic. The figure also includes the corresponding region of the sperm and pygmy sperm whales, plus a portion of the mitochondrial control region of the cow, which was used as an outgroup in the parsimony analysis.

As evident in figure 1, there is a great overall similarity between the 11 mysticete sequences. Differences between these sequences occur primarily in a region corresponding to positions 1-171 of the alignment in figure 1. Length differences are particularly conspicuous between positions 118 and 170. Starting with position 172 of the alignment, there is great similarity between different mysticetes, with respect to both sequence and length, although the bowhead, the right whale, and the pygmy right whale all lack an 11-bp region at around position 290 of the alignment. The two sperm whale sequences differ from those of the mysticetes, and the difference between the two species is large relative to that between different mysticetes.

The occurrence of the motifs TACAT and the complementary intranastrand ATGTA in the control regions of mammalian mtDNA has been discussed by Saccone et al. (1991). Both motifs occur in the mysticete and sperm whale sequences, but the possible secondary structures of these regions differ radically between the two groups.

The central domain of the mitochondrial control region was used by Saccone et al. (1991) to construct a phylogenetic tree including man, common chimpanzee, pygmy chimpanzee, Commerson dolphin, cow, rabbit, mouse, and rat. According to Saccone
et al. (1991), this portion of the mtDNA molecule can provide reliable dating of evolutionary separations. The dating of the separation between the cow and the dolphin, \( \approx 27 \) MYBP, is, however, at variance with mtDNA comparisons between the fin whale and the cow (Árnason et al. 1991a), which placed their separation at \( \approx 55 \) MYBP. However, the figure \( \approx 55 \) MYBP may not represent the dating of the artiodactyl/cetacean separation, which may have occurred later, provided that the cetaceans evolved from an artiodactyl lineage other than that represented by the cow.

The central portion of the mysticete control regions is very similar and provides limited resolution. Therefore the phylogenetic comparison between the different mysticetes and the two sperm whales was based on the entire control region except the highly variable portion. The region underlying the comparison corresponds to position 172(TGTATAAA . . .)-end in figure 1. Bovine sequence (Anderson et al. 1982) was used as an outgroup. The pairwise differences between the 14 sequences are shown in table 1. The large values between the sperm whale and the remaining cetaceans, except the pygmy sperm whale, are greater than that between the cow and the same cetacean species. Within genus *Balaenoptera* the closest relationship was registered between the sei whale and the Bryde’s whale. The difference between these two species was only 1.7%. The corresponding figure for the blue and the fin whale was 3.4%, and that between the two minke whale specimens was 5.2%. Thus, the difference between the Atlantic and Antarctic minke whale specimens was distinctly greater than that between acknowledged species of the same genus.

Figure 2 shows a phylogram based on the sequences in figure 1, less their variable portion. Bootstrap values for 500 replications are also included. The heuristic and the branch-and-bound parsimony analyses yielded identical results. The tree shows that the primary separation among the mysticetes is between the two balaenids (the bowhead and the northern right whale) and the remaining species. In the latter group the pygmy right whale has a separate position. The distance values for this species are high, as is apparent in table 1. Among the rorquals (family *Balaenopteridae*), the two minke whale types are joined on a separate branch. Noteworthy is the similarity between the sequence of the gray whale (family *Eschrichtiidae*) and those of the rorquals. Within genus *Balaenoptera* the most closely related species pair, the sei whale and the Bryde’s whale, was identified in all bootstrap replications.

**Discussion**

Mysticete relationships were recently studied by using the taxonomic distribution of three unrelated DNA satellites (Árnason and Best 1991). The findings showed a clear evolutionary distinction between the bowhead whale and the right whale (family *Balaenidae*, sensu stricto) and the remaining eight acknowledged species. We suggested that the pygmy right whale should be removed from the family *Balaenidae*, sensu latu. The comparisons upheld the view (Gray 1874; Miller 1923; Barnes and McLeod 1984) that the pygmy right whale should constitute a separate family, the Neobalaenidae. The present analyses of the mitochondrial control region are consistent with this view. It is also supported by sequence comparisons of the heavy mysticete DNA satellite (Adegoke et al., accepted) and by sequence data of the common cetacean DNA satellite in all mysticete families (Árnason et al. 1992). Both studies showed that the Balaenidae, sensu stricto, constitute a clade separated from all other mysticetes, including the pygmy right whale. The studies also showed that the difference between genus *Balaenoptera* and the conframilial humpback whale (genus *Megaptera*) was similar to that between genus *Balaenoptera* and the gray whale (family *Eschrichtiidae*). Thus, the
Fig. 1.—Sequences of the mitochondrial control region (L-strand) of 11 mitocho
A bovine sequence, truncated at the 5' end, is also shown. Bmy = Balæna mysticetus, bowhead; Egl = Eubalaena glacialis, right whale; Bac = Balaenoptera acutorostrata, minke whale; Bbon = Balaenoptera bonaerensis, Antarctic minke whale; Bbo = Balaenoptera borealis, sei whale; Bed = Balaenoptera edeni, Bryde's whale; Bmu = Balaenoptera musculus, blue whale; Bpl = Balaenoptera physalus, fin whale; Mno = Megaptera novaeangliae, humpback whale; Ero = Eschrichtius robustus, gray whale; Cma = Caperea marginata, pygmy right whale; Kbr = Kogia breviceps, pygmy sperm whale; Fma = Physeter macrocephalus, sperm whale; and Bta = Bos taurus, cow. The sequence of Bbon is from Hoelzel et al. (1991), and that of the cow is from Anderson et al. (1982). Total lengths of cetacean sequences are shown at the end of each sequence. Indels are indicated by a dot (·). The bovine sequence marks the region underlying fig. 2 and table 1.
FIG. 1. (Continued)
Table 1
Pairwise Distances between Mysticetes, Sperm Whales, and Cow, in the Control Region of mtDNA

<table>
<thead>
<tr>
<th></th>
<th>MYSTICETES</th>
<th>Sperm Whale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bmy 1</td>
<td>Egl 2</td>
</tr>
<tr>
<td>1 Bmy</td>
<td>0.049</td>
<td>0.108</td>
</tr>
<tr>
<td>2 Egl</td>
<td>0.112</td>
<td>0.119</td>
</tr>
<tr>
<td>3 Bac</td>
<td>0.052</td>
<td>0.058</td>
</tr>
<tr>
<td>4 Bbon</td>
<td>0.052</td>
<td>0.067</td>
</tr>
<tr>
<td>5 Bbo</td>
<td>0.067</td>
<td>0.067</td>
</tr>
<tr>
<td>6 Bed</td>
<td>0.067</td>
<td>0.067</td>
</tr>
<tr>
<td>7 Bmu</td>
<td>0.034</td>
<td>0.046</td>
</tr>
<tr>
<td>8 Bph</td>
<td>0.041</td>
<td>0.048</td>
</tr>
<tr>
<td>9 Mno</td>
<td>0.057</td>
<td>0.057</td>
</tr>
<tr>
<td>10 Ero</td>
<td>0.123</td>
<td>0.123</td>
</tr>
<tr>
<td>11 Cma</td>
<td>0.134</td>
<td>0.134</td>
</tr>
<tr>
<td>12 Kbr</td>
<td>0.102</td>
<td>0.102</td>
</tr>
<tr>
<td>13 Pma</td>
<td>0.102</td>
<td>0.102</td>
</tr>
<tr>
<td>14 Bta</td>
<td>0.254</td>
<td>0.254</td>
</tr>
</tbody>
</table>

NOTE.—The values refer to the less-variable portion on the mtDNA control region, position 172 (TGTA . .) to the end, in the alignment in fig. 1. Data below the diagonal are no. of differences; and data above the diagonal are mean distances (adjusted for missing data). Designations of species are the same as in fig. 1.
Fig. 2.—Heuristic phylogram (PAUP; Swofford 1993) with branch-length values showing the relationship between the 11 mysticete and 2 sperm whale sequences of the stable portion of the mtDNA control region. A bovine sequence was used as an outgroup. Bootstrap values (500 replications, 50% majority-rule consensus) are shown underlined. Abbreviations are as in fig. 1. The branch lengths of the sperm whales are notably long. Among the mysticetes, the right whales (Bmy and Egl) separate from the remaining species, including the pygmy right whale (Cma). The close relationship between the gray whale (Ero) and the family Balaenopteridae (the rorquals) is noteworthy. The two minke whale sequences fall outside the clade including other rorquals and the gray whale.

previously identified close relationship between the Eschrichtiidae and the Balaenopteridae is in accord with the present mtDNA findings.

The molecular relationships within genus *Balaenoptera* have been dealt with in
considerable detail (Arnason et al. 1991b; Spilliaert et al. 1991; Wada and Numachi 1991; Arnason and Gullberg, accepted). In their alleloyme analysis, Wada and Numachi (1991) reported a very close relationship between the sei whale and the Bryde's whale. They also showed that the differences between the two species are considerably less than those between some populations of the minke whale.

Omura (1975), on the basis of morphological and osteological comparisons, and Rice (1977) proposed the recognition of three subspecies of the minke whale: *B. acutorostrata acutorostrata* (North Atlantic), *B. a. davidsoni* (North Pacific), and *B. a. bonaerensis* (Antarctic). Arnold et al. (1987) suggested the possibility that the Southern Hemisphere form should be given full species status, but these authors, like Stewart and Leatherwood (1985), emphasized that the final conclusion would depend on additional studies.

Comparison of the mitochondrial control region showed that the differences between the two minke whale types (the North-Atlantic *acutorostrata* and the Antarctic *bonaerensis*) are considerably greater than those between any other *Balaenoptera* species pairs (table 1). The differences between the two minke whale types are similar to those between the bowhead and the right whale or between the genus *Balaenoptera* and the humpback (genus *Megaptera*) or between family Balaenopteridae and family Eschrichtiidae. We therefore suggest that *B. a. bonaerensis* should be given full species status: *B. bonaerensis*, Antarctic minke whale. The definition of *B. a. davidsoni* will depend on additional analysis of that particular type.

The evolutionary separation of the blue whale and the fin whale has been estimated to have occurred \( \approx 5 \) MYBP (Arnason and Gullberg, accepted). In the stable portion of the mitochondrial control region the difference between the two species is 3.4%. The corresponding figure for the two minke whale specimens is 5.2%. Provided that there is a reasonable linearity in the evolution of this region, the evolutionary separation of the two minke whale types occurred \( \approx 7.5 \) MYBP. This figure is in accord with the datings by Wada and Numachi (1991).

The difference between the control regions of the mysticetes and those of the sperm whales is very large. The separate position of the sperm whales vis-à-vis other cetaceans has been documented both at the molecular level, by sequence analyses of the common cetacean DNA satellite (Grétarsdóttir and Arnason 1993), and at the chromosomal level (Arnason and Benirschke 1973; Arnason 1981a, 1981b). These studies suggest that the sperm whales have had a long separate evolution. In light of the great similarity between the mysticete mitochondrial control regions, the striking dissimilarity between the two sperm whale species may seem unexpected. However, the sperm whale and the pygmy sperm whale are systematically very distinct, leading Barnes et al. (1985) to suggest that they should have the status of different families: Physeteridae and Kogiidae, respectively.

Acknowledgments

We thank Dr. James G. Mead of the Smithsonian Institution, Washington, D.C., for tissue samples of the pygmy sperm whale. We also thank Dr. Walter M. Fitch for valuable comments on the manuscript. This work was supported by grants from the Swedish Natural Sciences Research Council and the Erik Philip-Sörensens Foundation.

LITERATURE CITED

ADEGOKE, J. A., Ü. ÁRNASON, and B. WIDEGREN. Sequence organization and evolution, in all extant whalebone whales, of a DNA satellite with terminal chromosome localization. Chromosoma (accepted).

ÁRNASON, Ú., and A. GULLBERG. The mtDNA sequence of the blue whale and a comparison between the mtDNAs of the blue and the fin whale, two species that can hybridize in nature. J. Mol. Evol. (accepted).
ÁRNASON, Ú., R. SPILLIAERT, Á. PÁLSĐÓTTIR, and A. ÁRNASON. 1991b. Molecular identification of hybrids between the two largest whale species, the blue (Balaenoptera musculus) and the fin whale (B. physalus). Hereditas 115:183–189.
MILLER, G. S. 1923. The telescoping of the cetacean skull. Smithsonian Miscellaneous Collections 76:1–72.


WESLEY M. BROWN, reviewing editor

Received May 27, 1992; revision received May 26, 1993

Accepted May 26, 1993